

BAB 6 PENGUJIAN DAN ANALISIS

Bab ini menjelaskan tentang pengujian yang dilakukan terhadap sistem penelitian yang sudah dibuat serta analisis terhadap hasil dari pengujian yang telah dilakukan.

6.1 Pengujian Variasi Rasio Seleksi Fitur

Pengujian seleksi fitur ini menguji hasil dari seleksi fitur data latih yang akan menjadi patokan perhitungan klasifikasi oleh data uji. Apakah dengan diberikan seleksi fitur pada analisis sentimen ini memberikan pengaruh yang lebih baik atau tidak sama sekali. Pengujian ini mencakup dua lingkup yaitu skenario pengujian dan analisis skenario pengujian.

6.1.1 Skenario Pengujian Variasi Rasio Seleksi Fitur

Pengujian ini dilakukan untuk mengetahui bagaimana pengaruh seleksi fitur dalam analisis sentimen. Pada implementasinya, seleksi fitur bisa diubah rasionya supaya bisa ditentukan berapa besar dimensinya. Sebagai contoh rasio 50 yang mempunyai arti 50% berarti dari kata-kata unik yang ada pada data latih hanya akan dipilih 50% kata atau setengahnya saja yang akan menjadi patokan untuk klasifikasi, kata-kata yang terpilih adalah kata yang mempunyai nilai *Query Expansion Ranking* terendah. Jika rasio yang diisi semakin besar, tentu kata-kata yang terpotong semakin kecil juga, jika parameter diisi 100 yang berarti 100% maka semua kata-kata akan diambil atau dengan kata lain tidak dilakukan seleksi fitur .

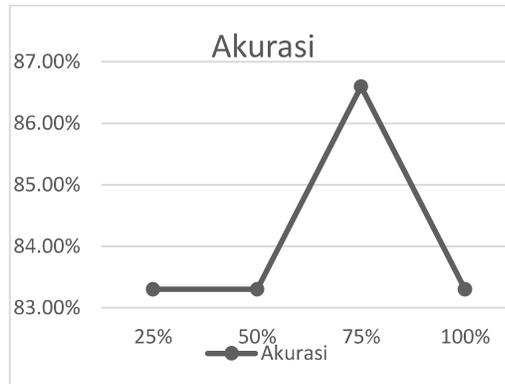
Skenario pengujian ini adalah dengan mengganti rasio seleksi fitur dengan 25, 50, 75, 100 menggunakan jumlah data latih dan data uji pada setiap pengujian yang sama, yaitu 200 data latih (100 data positif dan 100 data negatif) dan data uji sebanyak 30 dokumen. Semua percobaan variasi rasio seleksi fitur dilihat dan dicatat perubahan akurasi kemudian dianalisis. Untuk hasil dari pengujian ini dapat dilihat pada Tabel 6.1.

Tabel 6.1 Hasil pengujian seleksi fitur

No	Fitur	Jumlah Fitur	Akurasi
1	25%	472	83.3%
2	50%	944	83.3%
3	75%	1416	86.6%
4	100%	1888	83.6%

6.1.2 Analisis Pengujian Variasi Rasio Seleksi Fitur

Pengujian seleksi fitur dengan variasi rasio ini menghasilkan tingkat akurasi tertinggi pada seleksi fitur 75% dengan tingkat akurasi 86.6%. perubahan akurasi terhadap variasi rasio dalam bentuk grafik diilustrasikan dengan Gambar 6.1.



Gambar 6.1 Pengujian seleksi fitur

Dari Gambar 6.1 dapat dilihat bahwa akurasi terhasap variasi rasio seleksi fitur hamper sama yaitu pada 83.3 % pada rasio 25% dan 50%. Namun mengalami peningkatan pada rasio 75% sehingga mencapai akurasi 86.6% namun ketika rasio seleksi fitur dinaikkan lagi menjadi 100% akurasi kembali turun ke 83.3% seperti akurasi pada rasio seleksi fitur 25% dan 50%. Hal ini bisa terjadi karena pada rasio 25% dan 50% fitur-fitur yang dipotong jumlahnya sangat banyak sehingga beberapa fitur penting yang berpengaruh dalam proses klasifikasi hilang. Sedangkan pada rasio 100% dimana semua fitur digunakan membuat proses klasifikasi menjadi tidak efektif dan efisien karena dimensi fitur terlalu besar sehingga banyak fitur yang tidak mempunyai pengaruh terhadap proses klasifikasi ikut terproses.

Sebagai contoh, pada salah satu dokumen uji dengan kelas negatif terdapat kata “menarik” dan “mahal” yang berpengaruh pada kelas positif dan negatif. Namun rasio seleksi fitur 25% dan 50% tidak memilih kata “mahal” sehingga kelas menjadi kelas positif, sedangkan pada rasio 100% dimana semua fitur dipilih, fitur lain yang kurang penting seperti “tempat” juga ikutkan dalam proses klasifikasi sehingga membuat hasil klasifikasi menjadi positif. Sedangkan rasio seleksi fitur 75% berhasil memilih fitur penting yaitu “mahal” dan “menarik” serta tidak memilih fitur “tempat” sehingga hasil klasifikasi benar.